



**HAL**  
open science

# Foreground Detection via Robust Low Rank Matrix Factorization including Spatial Constraint with Iterative Reweighted Regression

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

► **To cite this version:**

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah. Foreground Detection via Robust Low Rank Matrix Factorization including Spatial Constraint with Iterative Reweighted Regression. International Conference on Pattern Recognition, ICPR 2012, Nov 2012, Tsukuba, Japan. pp.2805-2808. hal-00809470

**HAL Id: hal-00809470**

**<https://hal.science/hal-00809470>**

Submitted on 13 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Foreground Detection via Robust Low Rank Matrix Factorization including Spatial Constraint with Iterative Reweighted Regression

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

Laboratoire MIA (Mathematiques, Image et Applications) - University of La Rochelle  
ezahzah@univ-lr.fr

## Abstract

*Foreground detection is the first step in video surveillance system to detect moving objects. Robust Principal Components Analysis (RPCA) shows a nice framework to separate moving objects from the background. The background sequence is then modeled by a low rank subspace that can gradually change over time, while the moving foreground objects constitute the correlated sparse outliers. In this paper, we propose to use a low-rank matrix factorization with IRLS scheme (Iteratively reweighted least squares) and to address in the minimization process the spatial connexity of the pixels. Experimental results on the Wallflower and I2R datasets show the pertinence of the proposed approach.*

## 1. Introduction

The detection of moving objects is the basic low-level operations in video analysis. This detection is usually done using foreground detection. This basic operation consists of separating the moving objects called "foreground" from the static information called "background". Recent research on robust PCA shows qualitative visual results with the background variations approximately lying in a low dimension subspace, and the sparse part being the moving objects. First, Candes et al. [1] proposed a convex optimization problem to address the robust PCA problem. The observation matrix is assumed represented as:

$$A = L + S \quad (1)$$

where  $L$  is a low-rank matrix and  $S$  must be sparse matrix with a small fraction of nonzero entries. This research seeks to solve for  $L$  with the following optimization problem:

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 \quad \text{subj } A = L + S \quad (2)$$

where  $\|\cdot\|_*$  and  $\|\cdot\|_1$  are the nuclear norm (which is the  $L_1$  norm of singular values) and  $l_1$  norm, respectively, and  $\lambda > 0$  is an arbitrary balanced parameter. Under these minimal assumptions, this approach called Principal Component Pursuit (PCP) solution perfectly recovers the low-rank and the sparse matrices. Candes et al. [1] showed results on face images and background modeling that demonstrated encouraging performance. Several algorithms have been proposed for solving PCP such as Augmented Lagrangian Method [1], Low-Rank Representation [5] and SADAL[6]. However, PCP presents the following limitations: 1) The low-rank component are assumed to be exactly low-rank and the sparse component to be exactly sparse but the observations in real applications are often corrupted by noise affecting every entry of the data matrix; 2) The spatial connexity of the pixels are not addressed as the images are stacked in columns; 3) PCP is a batch algorithm. For the first limitation, Zhou et al. [11] proposed a stable PCP that guarantee stable and accurate recovery in the presence of entry-wise noise. Recently, Tang and Nehorai [9] proposed a PCP method via a decomposition that enforces the block sparsity of  $S$  and then address the second limitation. For the third limitation, He et al. [3] proposed an incremental Grasmannian RPCA method to update  $L$  and  $S$  when a new frame arrives.

In this paper, we propose a robust low-matrix factorization with IRLS scheme to address the second limitation. For a data matrix  $A$  containing the sequence, we assume that a part is approximately low-rank and product of two matrices, and a small part of this matrix is corrupted by the outliers. Furthermore, we directly introduced a spatial term in the  $l_1$  minimization to address the spatial connexity of the pixels. So, our contributions can be summarized as follows: 1) Addition of spatial constraint to minimization process, 2) IRLS alternating scheme for weighted the 2-parameters  $\|\cdot\|_{\alpha,\beta}$  for matrix low-rank decomposition. The rest of this paper is organized as follows. The Section 2 focus on IRLS method apply on vector regression problems. In

Section 3, we present a robust low-rank matrix factorization which allows us to detect foreground objects in dynamic backgrounds. In Section 4, we present comparison and evaluation versus the state-of-the-art methods and the Section 5 provides the conclusion.

## 2 $L_p$ Minimization with spatial constraint

In most applications, video surveillance data is assumed to be composed of background, foreground and noise. Regression task is a crucial part of the proposed decomposition algorithm and illustrate the main keys through a simple 1-D example shown in Fig. 1. We consider the following minimization problem (3), where  $A$  is a dictionary matrix (row order) and  $b$  is a row vector, the second term forces the error  $E$  to be a connexe shape, through the TV (Total Variation) of the error must be small, where the matrix  $\nabla_s$  is a spatial gradient.

$$\operatorname{argmin}_x \|Ax - b\|_\alpha + \lambda \|\nabla_s E\|_1, \quad E = |b - Ax| \quad (3)$$

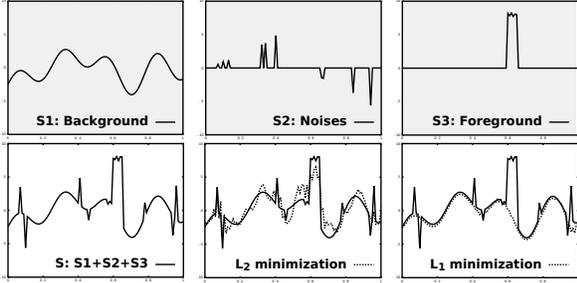


Figure 1. Schematic example for the 1-D case: We illustrate different fitting strategies of random dictionary basis (cosine function, wavelets, ...) on a composite signal  $S = S_1 + S_2 + S_3$ .

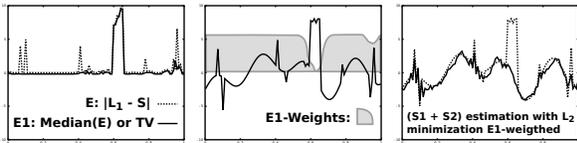


Figure 2. Overview of our background modeling strategy for the 1-D case with  $L_1$  minimization and spatial constraint in order to split Foreground and Background+Noises components.

The left part of the problem (3) is convex for  $\alpha > 1$  and the usual IRLS (Iteratively reweighted least squares) scheme for solve  $\operatorname{argmin}_x \|Ax - b\|_\alpha$  is given by

$$\begin{aligned} D^{(i)} &= \operatorname{diag}((\varepsilon + |b - Ax^{(i)}|)^{\alpha-2}) \\ x^{(i+1)} &= (A^t D^{(i)} A)^{-1} A^t D^{(i)} b \end{aligned} \quad (4)$$

It was proven that a suitable IRLS method is convergent for  $1 \leq \alpha < 3$  (cf. [2] and [8]). Since if the process is expressed with a residual formulation, we gain more numerical stability and let us to choose freely  $\alpha \in [1, \infty[$

with an adapted step size  $\lambda_{opt}$  on every iteration.

$$\begin{aligned} r^{(i)} &= b - Ax^{(i)} \\ D &= \operatorname{diag}((\varepsilon + |r^{(i)}|)^{\alpha-2}) \\ y^{(i)} &= (A' D A)^{-1} A' D r^{(i)} \\ x^{(i+1)} &= x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \end{aligned} \quad (5)$$

With a fixed  $\lambda_{opt}$ , we should choose  $\lambda_{opt} = \Lambda(\alpha)$ .

$$\Lambda(\alpha) = \begin{cases} 5/6 & \text{if } \alpha \leq 1 \\ -\frac{2}{3}\alpha + \frac{3}{2} & \text{if } 1 < \alpha < 1 + \frac{3}{4} \\ \frac{1}{\alpha-1} - 1 & \text{if } \alpha \geq 1 + \frac{3}{4} \end{cases} \quad (6)$$

Otherwise, the algorithm is twice iterative, where we try to get an optimal  $x$  and an optimal  $\lambda$  at each step.

$$\begin{aligned} c^{(i)} &= Ay^{(i)} \\ d^{(i)} &= b - A(x^{(i)} + y^{(i)}) \\ \operatorname{argmin}_{\lambda_{opt}} \|c^{(i)} \lambda - d^{(i)}\|_\alpha & \begin{cases} \lambda^{(0)} = \Lambda(\alpha) \\ s^{(k)} = d - \lambda^{(k)} c \\ E = \operatorname{diag}((\varepsilon + |s^{(k)}|)^{\alpha-2}) \\ z^{(k)} = \frac{c^t E s^{(k)}}{c^t E c} \\ \lambda^{(k+1)} = \lambda^{(k)} + (1 + \Lambda(\alpha)) z^{(k)} \end{cases} \end{aligned} \quad (7)$$

Only few iterations ( $\approx 10$ ) is enough for acceptable approximation of  $\lambda_{opt}$  of the  $\lambda^{(k)}$  sequence. Moreover, the convergence is usually improved by a Aitken process or an other acceleration technique. Note for case  $\alpha > 2$ , convergence is achieved when  $0 < 1 + \lambda < \frac{2}{\alpha-1}$ . Additionally, TV is particular case of the following problem:

$$\operatorname{argmin}_x \|Ax - b\|_\alpha + \lambda \|Cx - d\|_\beta \quad (8)$$

By derivation, the associated IRLS scheme is,

$$\begin{aligned} r_1 &= b - Ax^{(i)}, r_2 = d - Cx^{(i)}, e_1 = \varepsilon + |r_1|, e_2 = \varepsilon + |r_2| \\ D_1 &= (\sum e_1^\alpha)^{\frac{1}{\alpha}-1} \operatorname{diag}(e_1^{\alpha-2}), D_2 = \lambda (\sum e_2^\beta)^{\frac{1}{\beta}-1} \operatorname{diag}(e_2^{\beta-2}) \\ y^{(i)} &= (A' D_1 A + C' D_2 C)^{-1} (A' D_1 r_1 + A' D_2 r_2) \\ x^{(i+1)} &= x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \end{aligned} \quad (9)$$

The penalty term of (3) used for to increase connexity of the error requires to solve this minimization problem,

$$\operatorname{argmin}_x \|C|Ax - b|^\alpha - d\|_\beta \quad (10)$$

With  $\circ$  denotes the Hadamard product, the IRLS follow:

$$\begin{aligned} r_1 &= b - Ax^{(i)}, r_2 = d - C|Ax^{(i)} - b| \\ D &= \operatorname{diag}((C' |r_2|^{\beta-2} \circ r_2) \circ |r_1|^{\alpha-2}) \\ y^{(i)} &= (A' D A)^{-1} (A' D r_1), \quad x^{(i+1)} = x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \end{aligned} \quad (11)$$

More generally, we consider the following matrix regression problem  $\min_X \|AX - B\|_{\alpha, \beta, W}$  with two parameters norm  $(\alpha, \beta)$  and a weighted matrix  $(W)$ ,

$$\text{With } \|M_{ij}\|_{\alpha, \beta, W} = \left( \sum_{i=1}^n \sum_{j=1}^m W_{ij} |M_{ij}|^\beta \right)^{\frac{1}{\alpha}} \quad (12)$$

The problem is solved in the same manner on matrices with a reweighted regression strategy,

$$\begin{aligned} &\text{Until } X \text{ is stable, repeat on each } k\text{-columns} \\ \begin{cases} R & \leftarrow B - AX \\ S & \leftarrow \varepsilon + |R| \\ D_k & \leftarrow \operatorname{diag}(S_{ik}^{\beta-2} \circ (\sum_j (S_{ij}^\beta \circ W_{ij}))^{\frac{\alpha}{\beta}-1} \circ W_{ik})_k \\ X_{ik} & \leftarrow X_{ik} + (1 + \Lambda(\max(\alpha, \beta))) (A^t D_k A)^{-1} A^t D_k R_{ik} \end{cases} \end{aligned} \quad (13)$$

### 3 Foreground Detection via Robust Low-Rank Matrix Factorization

The training video sequence  $A \in \mathbb{R}^{n \times m}$  is store as a matrix with a particular structure. Columns are spatial frames and rows are values of a fixed pixel over time. For  $A = \{I_1, \dots, I_m\}$ ,  $I_j$  denotes a vectorized frame of  $n$  pixels at  $j$ -time with  $m$  is the number of frames.  $A_{y+hx,t}$  implies the pixel intensity at coordinate  $x, y, t$ . Then, the decomposition involves the following model:

$$A = L + S = BC + S \quad (14)$$

where  $B$  is a *low-rank* matrix corresponding to the background model plus noise and  $C$  allows to approximate  $L$  by linear combination.  $S$  is a *sparse* matrix which corresponds to the foreground component obtained by subtraction. The model involves the error reconstruction determined by the following constraints:

$$\min_{B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times m}} \|(A - BC) \circ W\|_{\alpha, \beta} + \mu \|BC\|_* \quad (15)$$

where  $\|\cdot\|_*$  denote the nuclear norm. The decomposition is fragmented into two parts. Firstly, we track 1-Rank decomposition since the first eigen-vector is strongly dominant in video surveillance.

$$\begin{aligned} R_1 &= A - B_1 C_1 \\ R &= A - B_1 C_1 - B_r C_r \end{aligned} \quad \left| \begin{array}{l} \min_{B_1, C_1} \|R_1\|_{1,1} \\ \min_{B_r, C_r} \|R \circ \phi(R_1)\|_{2,1 \rightarrow 0} \end{array} \right. \quad (16)$$

We use  $\|\cdot\|_{2,1 \rightarrow 0}$  instead of usual  $\|\cdot\|_{1,1}$  because it forces spatial homogeneous fitting. Besides  $\beta = (1 \rightarrow 0)$  means the  $\beta$  parameter decreases during iteration. First, we search a solution of the convex problem  $\|\cdot\|_{2,1}$ , then use the solution as an initial guess for non-convex problem  $\|\cdot\|_{2,(1-\varepsilon)}$ . Finally, we find a local minimum of  $\|\cdot\|_{2,0}$  and hope that is near of the global minimum of this problem. In the case where  $\alpha = \beta = 2$ , the decomposition is usually solved by a SVD (Singular Value Decomposition). Thus, our SVD algorithm can be seen as an iterative regression. The proposed scheme determines alternatively the optimal coefficients, it means searching  $C$  for  $B$  fixed and searching  $B$  for  $C$  fixed.

$$\begin{aligned} C^{(k+1)} &= (A^t A)^{-1} A^t B^{(k)} \\ \bar{C}^{(k+1)} &= C^{(k+1)} \sqrt{C^{t(k+1)} C^{(k+1)}}^{-1} \\ B^{(k+1)} &= (A^t A)^{-1} A^t \bar{C}^{(k+1)} \end{aligned} \quad (17)$$

Additionally, this alternating regression framework allows to associate a weighed matrix  $W$  which is entry-wise multiplied to the error.  $\min_{B,C} \|(A - BC) \circ W\|_{\alpha, \beta}$

The  $W$  mask is iteratively computed and aims to enforce the fit exclusively on guessed background region.

We define a function  $\phi$  that have two goals, smooth the error (like spatial median filtering) and transform the error for obtain a suitable weighted mask for regression.

$$W = \phi(|A - BC|) , \quad \phi(E) = e^{-\gamma TV(E)} \quad (18)$$

By including local penalty as a constraint in RPCA, this explicitly increases local coherence of the sparse component as filled/plain shapes (therefore moving object).

### 4 Experimental Results

We have compared the proposed approach with recent RPCA approaches: LBD [9], LRR [5], SADAL [6] and GRASTA [3] algorithms. The experiments were conducted qualitatively and quantitatively on the Wallflower dataset [10] I2R dataset [4]. The algorithms are implemented with Matlab.

#### 4.1 Wallflower and I2R dataset

The Wallflower dataset provided by Toyama et al. [10] consists of seven video sequences, with each sequence presenting one of the difficulties a practical task is likely to encounter. The images are  $160 \times 120$  pixels. The Fig. 3 and Fig. 4 show the qualitative results. For the quantitative evaluation, we used metrics based on the detection rate, the precision and the F-measure [7] computed between algorithm result and hand-segmented ground truth. Table 1 shows the results obtained for each algorithms.

Sequence	Frame	LBD	LRR	SADAL	GRASTA	Our method
bootstrap	00299	<b>70.18</b>	69.40	67.02	55.37	58.78
campus	01650	59.93	59.99	68.14	61.84	<b>75.07</b>
curtain	22772	<b>91.08</b>	88.34	91.01	89.88	90.73
escalator	02424	65.28	63.41	59.81	<b>68.31</b>	63.68
hall	02926	73.58	69.75	77.57	78.25	<b>79.26</b>
shoppingmall	01862	80.15	77.88	<b>82.80</b>	79.54	81.94
watersurface	01499	90.37	83.57	<b>92.01</b>	90.97	91.02
camouflage	00251	70.58	70.49	76.00	70.34	<b>82.27</b>
fore. aperture	00489	60.69	50.10	71.55	75.14	<b>75.19</b>
light switch	01865	57.74	36.76	<b>69.33</b>	28.35	58.49
moved objects	00985	0	0	0	0	0
time of day	01850	71.43	54.41	<b>80.84</b>	79.80	80.43
waving trees	00247	62.65	50.74	81.67	84.16	<b>84.70</b>

Table.1 F-measure for LBD [9], LRR[5], SADAL[6], GRASTA[3], our method (direct one-to-one correspondence with Fig. 4).

The I2R dataset provided by [4] consists of nine video sequences, which each sequence presenting dynamic backgrounds or illumination changes. The size of the images is  $176 \times 144$  pixels. For each sequence, the ground truth is provided for 20 images. Among this dataset, we have chosen to show results on seven sequences (see Fig. 3 and Fig. 4).

#### 4.2 Discussion

The proposed approach outperforms the other algorithms on five sequences over the Wallflower and I2R datasets and it is followed by SADAL. For the



Figure 3. From top to bottom: Wallflower and I2R datasets. From left to right (split in two columns): The original image, the background model and their difference.

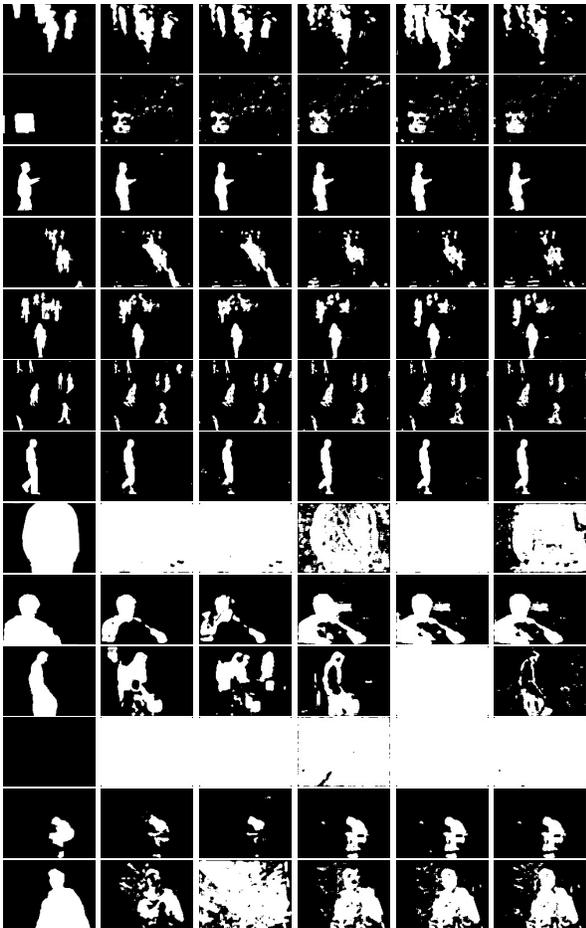


Figure 4. Foreground detection masks on the Wallflower and I2R datasets. From left to right: Ground Truth, LBD, LRR, SADAL, GRASTA, our method.

other sequences, the algorithm is in the second place or the results are still acceptable.

## 5 Conclusion

In this paper, we have presented a robust matrix factorization for foreground detection. This method is conceptually simple, easy to implement and efficient. Furthermore, experiments on video surveillance datasets show that this approach is more robust than recent RPCA approaches in the presence of dynamic backgrounds and illumination changes. Further research consists in developing an incremental version to update the model at every frame and to achieve real-time requirements.

## References

- [1] E. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *International Journal of ACM*, 58(3), May 2011.
- [2] I. Daubechies, R. Devore, M. Fornasier, and C. S. Guntkr. Iteratively reweighted least squares minimization for sparse recovery. *Comm. Pure Appl. Math*, 2008.
- [3] J. He, L. Balzano, and A. Sztam. Incremental gradient on the grassmannian for online foreground and background separation in subsampled video. *Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2012.
- [4] L. Li, W. Huang, I. Gu, and Q. Tian. Statistical modeling of complex backgrounds for foreground object detection. *IEEE Transaction on Image Processing*, pages 1459–1472, 2004.
- [5] Z. Lin, R. Liu, and Z. Su. Linearized alternating direction method with adaptive penalty for low-rank representation. *NIPS 2011*, Dec. 2011.
- [6] S. Ma. Algorithms for sparse and low-rank optimization: Convergence, complexity and applications. *Thesis*, 2011.
- [7] L. Maddalena and A. Petrosino. A fuzzy spatial coherence-based approach to background foreground separation for moving object detection. *Neural Computing and Applications*, pages 1–8, 2010.
- [8] M. R. Osborne. Finite algorithms in optimization and data analysis. *John Wiley & Sons*, 1985.
- [9] G. Tang and A. Nehorai. Robust principal component analysis based on low-rank and block-sparse matrix decomposition. *Annual Conference on Information Sciences and Systems, CISS 2011*.
- [10] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and practice of background maintenance. *ICCV 1999*, pages 255–261, September 1999.
- [11] J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao. Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization. *NIPS 2009*, December 2009.